Hybrid Stereo Camera

An IBR Approach for Synthesis of Very High Resolution Stereo Image Sequences

Harpreet S. Sawhney  Yanlin Guo  Sean Adkins
Keith Hanna  Rakesh Kumar  Samuel Zhou

http://www.sarnoff.com/search/tech_papers/hybrid/index.asp
Motivation

Extremely High Visual Quality

Stereo Creation & Projection
4K-8K digital resolution per eye
World’s Largest Film Format

IMAX 3D Content

CG Animations
Live Action
Mixed CG & Live Action
Limitations on IMAX 3D Content Creation

Live Action Content

- Camera is very large.
- Requires two strips of large format film.
- Size of camera and cost of film limits production.

CG Content

- 6-14 hours rendering time per frame!
Solution: Hybrid Stereo Camera

- Digital Camera taking lens
- Digital Camera
- Camera Electronics
- Removable Disk Array
- Film Camera taking lens
- Film Camera and Film Magazine
Goals

Expand the possibilities for 3D Cinematography:

Can Computer Vision & IBR deliver High Quality?

With reduced cost & time?

Explore an Analysis-Test-Synthesis Framework for Image-based Modeling & Rendering
Hybrid Stereo Camera

... pure upsampling is not an option ...

INPUT

OUTPUT

1:16

Left Eye (1.5K) Right Eye (6K) Left Eye (6K) Right Eye (6K)
Live Action Sequence
Live Action: Hybrid Input
Synthesized Output
How can the Hybrid Camera be Realized?

Render the High-Res content into the coordinate system of the Low-Res Frame!

Left Eye

Right Eye

t-2 t-1 t t+1 t+2
Approach
Convergence of Computer Vision & IBR

- Compute stereo disparities at lo-res.
- Compute motion (Optical Flow) at lo-res.
- Compute quality map at lo-res.
- Synthesize hi-res frame.
- Fill-in and color correct mis-matched pixels.
- Temporal de-scintillation.
Approach

Convergence of Computer Vision & IBR

• Compute stereo disparities at lo-res.

• Compute motion (Optical Flow) at lo-res.

• Compute quality map at lo-res.

• Synthesize hi-res frame.

• Fill-in and color correct mis-matched pixels.

• Temporal de-scintillation.
Establishing Stereo/Motion Correspondences

Analysis

Left Eye
Time t

Right Eye
Time t, t’

Synthesis
Correspondences by Coarse-to-fine Model-based Image Alignment: A Primer

\[
\min_{\Theta} \sum_p (I_1(p) - I_2(p + u(p;\Theta)))^2
\]
Correspondences by Coarse-to-fine Model-based Image Alignment: A Primer
3D Stereo Correspondence

Original Left-Right Pair
3D Stereo Correspondence

Disparity/Correspondence Map
3D Stereo Correspondence

Original Left & Disparity-warped Right Frame
Quality of Alignment Map

Associate a \([0,1]\) value at each pixel
Aggregate Quality Map at Lo-Res
Using Stereo-Motion Synthesis

Warped frame t-1 → Warped frame t → Warped frame t+1

Original Lo-res → Synthesized Lo-res → Aggregate Q-map
Filling-in Mismatched Pixels at Hi-res

Hi-res Q-map

q > thresh

Y

N

Hi-res Synthesized frame

Up-resed frame
Filling-in Mismatched Pixels at Hi-Res

Sample Result
Color Correction

Use Matched Pixels to solve for a Color Model:
\[ \text{Color}(p') = A \times \text{Color}(p) + b \]

Apply Model to Mismatched Pixels

Synthesized Hi-res frame

Up-resed frame

Hi-res Q-map
Color Correction

Sample Result
Quantitative Validation

### Misaligned Pixels

<table>
<thead>
<tr>
<th>Stage of Synthesis</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>11.3</td>
</tr>
<tr>
<td>2</td>
<td>8.7</td>
</tr>
<tr>
<td>3</td>
<td>3.2</td>
</tr>
<tr>
<td>4</td>
<td>1.3</td>
</tr>
</tbody>
</table>

#### Live Action

#### CG Animation

<table>
<thead>
<tr>
<th>Stage of Synthesis</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>17.1</td>
</tr>
<tr>
<td>2</td>
<td>8.6</td>
</tr>
<tr>
<td>3</td>
<td>1.4</td>
</tr>
</tbody>
</table>
Synthesis vs. Up-resing: Live Action
Synthesis vs. Up-resing: CG Animation
Computational Time

• Research Code: Currently about 45 mins. per 4K frame on an SGI 350 Mhz Octane.

• Optimizations can easily reduce the time to about 4-5 mins.

Potential cost/time reduction for a 45 min. feature

180 CPUs / 6 months → 30 CPUs / 2.5 months
Generalizations

Key Idea: IBMR can exploit the availability of lower resolution or other similar data for high quality rendering.
• Applied an Analysis-Test-Synthesis Framework to high quality stereo synthesis.

• Initial validation of quality of synthesis is very encouraging.

• Potential for new research and applications based on generalizations of the framework.
Acknowledgements

• Ed Lepieszo & Carol Harrison, IMAX
  - Help with demos and frame synthesis.

• Vince Paragano & Doug Corliss, Sarnoff
  - Software and systems support.

• Spans & Partner Inc., and IMAX
  - CG and Live Action Stereo Sequences.
Related Work

• Perception of mismatched stereo pairs
  - Julesz'71, Perkins'92, Stelmach et al.'00

• Multi-resolution 3D/image sequence analysis
  - Bergen et al.'92, Hanna et al.'93

• Alignment quality measures
  - Irani et al.'94, Szeliski'99, Lubin'92
Temporal De-scintillation

Problem

• Synthesis is uncorrelated over time.
• Produces temporal scintillation.

Solution

• Smooth quality maps over time before using them for compositing.
Outstanding Issues

- Evaluation of JND based quality maps.
- Adaptive combination of stereo and motion frames.
- Issues related to real hybrid camera design.
Motivation

Extremely High Visual Quality
Stereo Creation & Projection
4K-8K digital resolution per eye
World’s Largest Film Format

IMAX 3D Content
CG Animations
Live Action
Mixed CG & Live Action

IMAX®
Goals

Expand the possibilities for 3D Cinematography:

Can Computer Vision & IBR deliver High Quality?

With reduced cost & time?

Explore an Analysis-Test-Synthesis Framework for
Image-based Modeling & Rendering
Establishing Stereo/Motion Correspondences

Analysis

Left Eye
Time t

Right Eye
Time t, t’

Synthesis
Computational Time

- Research Code: Currently about 45 mins. per 4K frame on an SGI 350 Mhz Octane.
- Optimizations can easily reduce the time to about 4-5 mins.

Potential cost/time reduction for a 45 min. feature

- 180 CPUs / 6 months
- 30 CPUs / 2.5 months